

BLOOM'UN PROGRAMIN ÖĞELERİNE DAYALI DEĞERLENDİRME MODELİ BAĞLAMINDA İNGİLİZCE ÖĞRETİM PROGRAMININ ÇOKLU ANALİZİ

Yrd. Doç. Dr. Veli Batdı
Kilis 7 Aralık Üniversitesi
veb_27@hotmail.com

Özet

Bu araştırmanın amacı İlköğretim 5., 6., 7. ve 8. Sınıf İngilizce Öğretim Programını (İÖP), Programın Öğelerine Dayalı Değerlendirme Modeli (PÖDDEM) bağlamında çoklu analizle incelemektir. Nitel ve nicel boyutlar içeren çoklu analiz bir konuya ilişkin mevcut durumu, farklı programlarla analiz etmeyi içermektedir. Bu bağlamda nicel boyut kapsamında Rasch Modeliyle SPSS programı kullanılarak betimsel analiz; nitel bağlamda ise Maxqda 11 programıyla durum çalışmasına uygun tematik analiz yapılmıştır. 2016-2017 akademik yılında Elazığ merkezdeki ortaokul İngilizce öğretmenlerinden oluşan araştırma evreninden randomize yolla belirlenen 42 öğretmen Rasch modeli; 78'i SPSS analizleri ve maximum çeşitlilik örnekleme göre belirlenen 27'si tematik analizler için çalışma grubuna seçilmiştir. Rasch ve SPSS sonuçlarından 7. Sın. İÖP'ün en nitelikli; 5. Sın. en düşük nitelikli program olduğu anlaşılmıştır.

Anahtar Sözcükler: Programın öğelerine dayalı değerlendirme modeli, İngilizce öğretim programı, çoklu analiz.

MULTI-ANALYSIS OF ENGLISH CURRICULUM THROUGH BLOOM'S EVALUATION MODEL BASED ON PROGRAM ITEMS

Abstract

The aim of this study is to examine the 5th, 6th, 7th and 8th grades English Curriculum through an evaluation model based on program items considering multi-analysis. Including both quantitative and qualitative aspects, multi-analysis consists of a variety of statistical programs in order to examine a certain subject matter. As for descriptive analysis, Rasch evaluation model and SPSS program are used in the quantitative aspect while for thematic analysis Maxqda in the qualitative aspect. The working group consists of English teachers working at secondary schools in Elazığ during 2016 and 2017 academic years, 42 of whom are selected for Rasch and 78 are for SPSS analyses randomly. Also, 27 of these teachers are selected for thematic analysis through maximum variation sampling. As a result of Rasch and SPSS analyses, it is recorded that the 7th grade English Curriculum is the most qualified while the 5th one is the least.

Keywords: Evaluation model based on program items, English curriculum, multi-analysis.

GİRİŞ

Öğretim programları, kaliteli bir eğitim sisteminin yerleşmesi amacıyla her ders için ayrı ayrı geliştirilen ve belli kazanımlara sahip geniş kapsamlı bir plan olarak nitelendirilebilir. Milli eğitimin temel politikaları bağlamında geliştirilen öğretim programları öğretim etkinliklerinin neler olacağı, öğrenenlere nasıl ve ne zaman ulaştırılacağı noktalarında bilgi sağlamaktadır. Dolayısıyla öğretim programları, eğitim sistemimizin en önemli yönlerinden olan uygulama boyutunda neyin nasıl yapılacağına kılavuzluk etmektedir. Bu noktada programların işlevsel ve etkililik boyutunun sistemli olarak incelenmesi gerekli görülmektedir. Diğer bir ifadeyle, programlarının toplumsal ihtiyaç ve beklentilerine, öğrenenlerin ilgi, ihtiyaç ve özelliklerine, teknolojik gelişmelere, konu alanındaki değişim ve gelişmelere uygunluğunun değerlendirilmesi gerekmektedir.

Programı değerlendirme, öncelikle program geliştirme sürecine katkı sağlayacak nitelikte bilgi ve veri elde etme noktasında önemli görülmektedir (Stufflebeam ve Shinkfield, 1985). Burada program değerlendirme derken aslında programların işlevinden söz edilmektedir. Program değerlendirme yapıldıktan sonra programın iyi çalıştığı saptanmış ise, net bir bilgiye ulaşılmıştır. Bu noktada programın daha verimli ürün oluşturma sürecine katkı sunması amacıyla daha da geliştirilmesi sağlanabilir. Diğer yandan programın iyi çalışmadığı sonucuna ulaşırsa, tespit edilen bu durumun giderilmesi, eksikliğin tamamlanması amacıyla programlarda değişiklikler yapılması gerekebilir (Gökmenoğlu, 2014). Değerlendirme sonucu ulaşılan bilgiler program geliştirme politika ve programlarının iyileştirilmesinde kullanılmaktadır. Bu nedenle mevcut zaman ve kaynakların kullanımı ile ilgili verilecek kararlar son derece önemli görülmektedir (Wholey, Hatry ve Newcomer, 2010). Nitekim programın belirlenen zaman dilimi içinde yetiştirilmesi veya hangi kaynaklardan yararlanılması gerektiği ürün çıktısını doğrudan etkileyen durumlardır.

Program değerlendirme çalışmaları araştırma yapma ile benzer noktalara sahiptir (Fitzpatrick, Sanders ve Worthen, 2004). Her ikisinde de bilgiye ulaşmak için inceleme yapmak gerekmektedir. Ancak araştırma yapmada yeni bir bilgiye ulaşma ve onu keşfetme söz konusu iken program değerlendirmede mevcut bir durumun irdelenmesi, incelenmesi, bilgi toplanması ve yargıya varılması vurgulanmaktadır. Yargıya varma ise program değerlendirmenin en temel noktası olarak görülmektedir (Oliva, 1997). Program değerlendirme çalışmalarında en dikkat edilmesi gereken hususlardan biri yapılan işlemlerin sistematik, geçerli ve güvenilir olmasıdır. Bu şekilde yürütülen çalışmaların ardından kaliteli ve nitelikli ürünler elde edilebilir. Program değerlendirme çalışmalarının sistematik bir şekilde yürütülebilmesi amacıyla farklı ülkelerde farklı birçok program değerlendirme modeli geliştirilmiştir (Yüksel, 2010). Program değerlendirme modellerinin esas aldıkları felsefe ve yaklaşımlar da birbirinden farklılık göstermektedir. Amaçların belirlenmesine ve belirlenen amaçların sonuçlarına değinen amaca odaklı değerlendirme; yöneticilere bilgi sağlamaya dayalı yönetim odaklı değerlendirme (Worthen, Sanders ve Fitzpatrick, 1997); bağımsız değerlendiricilerin geliştirdiği eğitsel ürün ve hizmetlerin değerlendirilmesine dayalı tüketici odaklı değerlendirme modeli (Fitzpatrick ve diğ., 2004) bu kapsamda ileri sürülebilecek program değerlendirme modelleridir. Diğer yandan uzman odaklı; katılımcı odaklı ve rakip odaklı modeller de amaç ve kaynak açısından birbirinden farklı değerlendirme modelleri olarak göze çarpmaktadır (Gredler, 1996). Bütün bunların yanı sıra Ertürk (1998) program değerlendirme yaklaşımlarını altı ana grupta toplayarak bunları program tasarısı, ortam, başarı, erişimi, öğrenme ve ürün odaklı değerlendirme yaklaşımları olarak sıralamıştır. Ancak farklı yaklaşımları baz alarak geliştirilen program değerlendirme modellerinden denenmekte ve uygulanmakta olanlar Tyler'in Hedefe Dayalı Değerlendirme Modeli, Metseffel ve Michael Modeli, Provus'un Farklar Yaklaşımı Modeli, Stake'in Uygunluk-Olasılık Modeli, Stufflebeam'in Bağlam, Girdi, Süreç, Ürün (CIPP) Modeli, Stake'in İhtiyaca Cevap Vermeye Dayalı Değerlendirme Modeli, Eisner'in Eğitsel Uzmanlık/Eleştiri Modeli, Saylor, Alexander ve Lewis Modeli ve Bloom'un Programın Öğelerine Dayalı Değerlendirme Modeli olarak belirtilebilir.

Mevcut araştırmada Programın Öğelerine Dayalı Değerlendirme Modeli bağlamında programın değerlendirilmesi amaçlanmıştır. Bilindiği gibi programın temel öğeleri olan hedefler, içerik, öğrenme-öğretme süreçleri ve değerlendirme (Demirel, 1992) arasındaki ilişki dinamik ve birindeki küçük bir değişiklik diğerlerini de etkilemektedir. Dolayısıyla hedeflerin doğru belirlenerek öğrencilere kazandırılmaya çalışılması, içeriğin hedefler doğrultusunda belirlenmesi, öğrenme durumlarının hedeflere götüreceği biçimde düzenlenmesi ve değerlendirmede hedeflere ulaşıp ulaşılmadığının ölçülmesi tutarlı bir eğitim programının elde edilmesi için bir zorunluluktur (Bümen, 2006). Bu nedenle programın dört temel öğesi bağlamında yapılacak değerlendirmenin programların niteliği açısından önemli ve gerekli olduğu düşüncesiyle araştırmada Programın Öğelerine Dayalı Değerlendirme Modelinin dikkate alınması uygun görülmüştür.

Araştırmanın Amacı ve Önemi

Bu araştırmada, ilköğretim 5., 6., 7. ve 8. Sınıf İngilizce Öğretim Programının (İÖP) Bloom'un "programın öğelerine dayalı değerlendirme modeli (PÖDDEM)" bağlamında hedef, içerik, öğrenme durumları ve değerlendirme alt boyutları çerçevesinde Rasch, SPSS ve Maxqda programlarıyla değerlendirilmesi amaçlanmıştır. Ayrıca her analiz ile ulaşılan bulguların birbiriyle karşılaştırılarak yorumlanması düşünülmüştür. Program değerlendirme çalışmalarının bir ülkenin eğitim açısından gelişimine olan katkısı küçümsenemez niteliktedir. Nitekim yapılan değerlendirmelerde farklı örneklem gruplarından (müfettiş, yönetici, öğretmen, öğrenci, veli, vb.) alınan görüşlerin programların artı ve eksi yönlerini ortaya çıkararak sağlam dönüşlere ulaşılmasına etki ettiği belirtilebilir. Programların işlevsel boyutu, etkililiği ve sınırlılığı belirlenerek buna yönelik

tedbirler alınabilir ve farklı düzenlemelere gidilerek iyileştirmeler yapılabilir. Bu şekilde öğrenme sonucunu olumlu yönde etkileyecek gelişmelere yer verilmiş olur. Alanyazın taraması yapıldığında ülke çapında ve uluslararası çapta program değerlendirmeye yönelik farklı çalışmalara rastlanmıştır (Adamson ve Prion, 2015; Anderson, 2005; Çeliköz, 2004; Hubbard ve Bauer, 1993; Kim, 2012; Özdemir, 2009; Pekdağ ve Erol, 2013). Ancak özellikle PÖDDEM'e ilişkin literatürde çok fazla çalışmaya rastlanmadığı görülerek mevcut araştırmanın bu yönde yürütülmesi kararlaştırılmıştır. Diğer yandan alanyazında ulaşılan çalışmaların genel anlamda nitel veya nicel şeklinde tek boyutlu yürütüldüğü anlaşılmıştır. Mevcut çalışmamızın birden fazla analiz programını içeren çoklu analizle yürütülerek bu anlamda da fark yarattığı ve alana katkı sunduğu belirtilebilir. Bu noktada 'Programa Dayalı İnceleme ve Değerlendirme (Program-Based Review and Assessment)' adlı bir kitapta etkili program değerlendirme çalışmalarının çok boyutlu ve çok yönlü (multi-faceted) olmasının gerekliliği vurgulanarak (Stassen, Doherty ve Poe, 2001) bu durum mevcut çalışmamıza kaynaklık etmiştir.

Araştırmada kullanılan *çoklu analiz*, farklı kaynaklardan elde edilen verilerin farklı analiz programlarıyla çözümlenerek daha kapsamlı, güvenilir ve detaylı sonuçlara ulaşmasını sağlayan bir süreci içermektedir. Veri kaynaklarındaki çeşitlilik ve zenginlik olaylara farklı bakış açılarıyla bakarak araştırmanın niteliğini arttırmaktadır (Marvasti, 2004; Patton, 2002). Bu noktada mevcut çalışmada Rasch, SPSS ve Maxqda programları kullanılmıştır. Çoklu analiz kapsamında farklı analizler kullanılmasının amaç ve gerekçeleri bulunmakta ve bunlar EK-1'de detaylı bir şekilde açıklanmaktadır. Mevcut çalışmada İÖP'ün Rasch analizi ile incelenmesi sonucunda dört ayrı programın (5., 6., 7. ve 8. Sınıf) dört farklı faktör açısından (hedef, içerik, öğrenme durumları ve değerlendirme) değerlendirilmesine yer verilmiştir. Diğer bir ifadeyle, çoklu ürün değerlendirmesi yapılarak katılımcıların görüşleri bütüncül şekilde görülebilmektedir. Ancak SPSS Programı tek ürün değerlendirme odaklı olduğundan daha spesifik değerlendirmeler yapmaktadır. Sınıflar bazında genel aritmetik ortalama değerlerine ise Rasch ile ulaşılamadığından, bu aşamada SPSS programı işe koşulmuştur. Diğer yandan Rasch analizinde katılımcıların en yüksek (cömert) ve en düşük (katı) katılım düzeyleri belirlenebilmektedir. Cinsiyet gözetmeksizin yapılan bu analize katkı sağlama noktasında, SPSS programı kullanılarak anlamlı farklılığın hangi cinsiyet lehine olduğu da belirlenebilmektedir. Dolayısıyla analizlerin birbirleriyle karşılaştırıldığında, eksik görülen noktaların tamamlanması veya birinde ulaşılamayan bir bilginin diğeriyle ulaşabilir kılınması amacıyla çoklu analiz yapılmasına gerek duyulmuştur. Böylece eksik görülen bilgilerin farklı programlarla tamamlanması ve detaylı analizler yapılarak araştırmanın her yönüyle açıklık ve netlik kazanması sağlanabilir.

Özetle İÖP'ün PÖDDEM bağlamında çoklu ürün (programlar) değerlendirmesini yapmak ve katılımcı yanlılığını görmek Rasch ölçme modeliyle; katılımcıların ilgili ürünlere ilişkin katılım düzeylerini belirlemek ve cinsiyet değişkeni açısından yorum yapmak SPSS programıyla; programların PÖDDEM bağlamında katılımcı görüşlerine başvurularak incelenmesi ve model oluşturulması ise Maxqda Programıyla gerçekleştirilmiştir. Dolayısıyla İÖP'ü hedef, içerik, öğrenme durumları ve değerlendirme öğeleri bağlamında incelemek temel amacı doğrultusunda şu alt amaçlar belirlenmiştir:

İÖP'ün PÖDDEM'e göre;

- *Rasch Ölçme Modeli kapsamında*, katılımcı görüşlerinin genel analizi, jürilerin katılık/cömertlikleri, madde güçlük analizleri,
- *SPSS 17 programı doğrultusunda*, katılımcı görüşlerinin cinsiyet değişkeni açısından değerlendirilmesi ve İÖP öğelerinin aritmetik ortalamasının karşılaştırılması,
- *Maxqda 11 programı kapsamında*, hedef, içerik, öğrenme durumları ve değerlendirme öğeleri kapsamında belirtilen görüşlerin incelenmesi amaçlanmıştır.

YÖNTEM

Bu araştırma, hem nicel hem de nitel yöntemlerin birlikte kullanılmasını içeren çoklu analiz bağlamında yürütülmüştür. Bu bağlamda araştırmada, Rasch ölçme modeli, SPSS 17 ve Maxqda 11 Programları bir arada kullanılarak farklı analiz sonuçlarına gidilmiştir. Bu süreçte Tablo 1'de detaylandırıldığı gibi bir araştırma konusu çalışılırken, farklı programlar kullanılarak yapılan analizlerle daha eksiksiz ve bütüncü bilgilere ulaşılması amaçlanmıştır. Bu noktada her analizdeki mevcut olan sınırlı yön, diğer bir analiz türüyle tamamlanarak kapsamlı sonuçlara ulaşılmaya çalışılmıştır.

Araştırma Türü ve Deseni

Araştırmanın çoklu analiz kapsamında yer alan nicel boyutta, Rasch Ölçme Modeli ve SPSS programından yararlanılarak betimsel analize dayalı işlemler yürütülmüştür. Rash ölçme modeli bireyin belirli bir özelliğe ilişkin performansının verdiği tepkiyle öngörülmesini sağlamakta ve geçerli ölçümlerin gelişimi, değişimi ve takibine yönelik kayda değer veriler sunmaktadır (Boone ve Scantlebury, 2006). Betimsel analiz, verilerin daha anlaşılır biçimde düzenlenmesini içeren (Zikmund, 2003) veya bilimsel alanda gözlemi kaydetme, olaylar arasındaki ilişkiyi saptama ve genellemelere ulaşma şeklindeki bilimin betimleme yönünü kapsayan bir süreçtir (Yıldırım ve Şimşek, 2006). Nitel boyutta ise Maxqda programı ile durum çalışmasına uygun tematik bir inceleme yapılmıştır. Nitel bağlamda izlenen durum çalışması ise geniş çapta ve derinlemesine bir olay veya durumun gerçek ortamda incelenmesini içeren ve birden fazla veri kaynağının kullanıldığı (Yin, 1994) sınırlı alanda veya evrende yürütülen durumların detaylı incelendiği (Zainal, 2007) bir araştırma türü olarak belirtilebilir. Bu bağlamda İÖP'ün PÖDDEM'e göre analiz edilerek farklı analiz sonuçlarının karşılaştırılması bu araştırmanın temel amacını oluşturmuştur.

Çalışma Grubu

Çalışmanın evrenini Elazığ il merkezinde 2016-2017 akademik yılında Elazığ merkezdeki ortaokul İngilizce öğretmenleri oluşturmaktadır. Örneklemini ise bu evrenden seçilen ve 5., 6., 7., ve 8. sınıf İngilizce derslerine giren öğretmenler oluşturmuştur. Bu bağlamda araştırma evreninden randomize yolla seçilen 42 İngilizce öğretmeni Rasch modeli; 78 öğretmen SPSS analizi ve maksimum çeşitlilik örnekleme uygun seçilen 27 öğretmen ise tematik analizler için çalışma grubuna alınmıştır.

Araştırmanın Rasch ve SPSS verilerine ulaşmak amacıyla araştırmacı tarafından ilgili literatür tarandıktan sonra uzmanların (Eğitim Bilimleri Bölümünden 2 Profesör, 2 Doçent, 3 Yrd. Doçent, 1 Ölçme ve Değerlendirme uzmanı, ortaokulda görev yapan 10 İngilizce ve 4 Türkçe öğretmeni) görüşlerine başvurularak hazırlanan "5.-8. Sınıflar İngilizce Öğretim Programını Değerlendirme Formu" kullanılmıştır. Hazırlanan formdaki maddelerin Kapsam Geçerlik Oranı (KGO) ve Kapsam Geçerlik İndeksi (KGİ) değerleri hesaplanarak $KGİ > KGO / 0.82 > 0.56$ değerleri elde edilmiştir. Bu değerler anketin güvenilir olduğunu göstermiştir. 24 maddeden oluşan form beşli likert türünde (Hiç:1.00-1.80; Nadiren:1.81-2.60, Bazen: 2.61-3.40, çoğunlukla:3.41-4.20 ve Daima: 4.21-5.00) ve 4 faktörden (Hedef, İçerik, Öğrenme durumları, Değerlendirme) oluşmuştur. Form faktörlerinin iç tutarlılık değerleri .68 ile .90 arasındadır. Bu değerlerin +1.00'e yakın olması ölçeğin tutarlılık ve güvenilirliğinin yüksek olduğu şeklinde yorumlanabilir (Özdamar, 2010). SPSS analizleri yapılırken ortalama, standart sapma ve ikili karşılaştırmalarda t-testi kullanılmıştır. Araştırmada .05 değeri anlamlılık düzeyi olarak dikkate alınmıştır. Nitel kapsamda araştırmacı tarafından ilgili literatür taraması ve uzman görüşlerine dayalı geliştirilen bir görüşme formu ile veri toplanması sağlanmıştır. İlgili formda, 5., 6., 7. ve 8. Sınıf İÖP'e ilişkin İngilizce öğretmenlerinin görüşleri PÖDDEM'e dayalı hedef, içerik, öğrenme durumları ve değerlendirme faktörleri çerçevesinde alınmıştır. 27 öğretmenin görüşlerinin alınmasından sonra ulaşılan nitel veriler farklı tema ve kodlar şeklinde model haline getirilmiştir.

Farklı analiz türlerinin kullanıldığı mevcut çalışmamızda ulaşılan tüm veriler uygun şekilde ve sırada yorumlanarak sunulmuştur. Hem nicel hem de nitel yöntemin bir arada kullanıldığı (Bamberger, 2012) çoklu analiz ile ulaşılan veriler, İÖP'ün PÖDDEM'e dayalı değerlendirilmesi bağlamında genel bir bakış sunulması üzerine ayrı ayrı analiz edilmiştir. Araştırmanın nicel boyutu, Rasch ölçme modeliyle yapılmıştır. Buna göre araştırmanın üç yüzeyi bulunmaktadır. Birincisi İngilizce öğretmenlerinden oluşan 42 adet jüri; ikincisi ölçme aracında bulunan 24 madde ve üçüncüsü 5., 6., 7. ve 8. Sınıflar İÖP olarak belirtilebilir. Diğer yandan araştırmanın aritmetik ortalama, standart sapma ve faktörlere ilişkin analizlerini içeren nicel boyutu SPSS Programı ile yapılmıştır. *Nitel analiz kapsamında*, İÖP'ün PÖDDEM'e dayalı değerlendirilmesiyle ulaşılan görüşler Maxqda 11 programıyla analiz edilerek içerik analiziyle çözümlenmiştir. Bu aşama, araştırmaya zengin ve ayrıntılı veri kaynağı sağlayan ve esnek bir boyut kazandıran tematik analiz içermiştir (Braun ve Clarke, 2006). Tematik boyutta kullanılan görüşme formları, kodlayıcıların birbirinden bağımsız bir şekilde incelemesi sonucunda tema ve kodlar şeklinde oluşturulmuştur. Hedef, içerik, öğrenme durumları ve değerlendirme biçiminde oluşturulan her temanın kodları Maxqda programına yüklenerek model oluşturulmuştur. Ayrıca kodlayıcılar arasındaki güvenilirlik hesaplaması yapılarak Kappa uyum değerleri bulunmuştur. .81 ile .90 arasında bulunan bu değerler Viera ve Garrett'a (2005) göre çok iyi düzeyde uyum şeklinde yorumlanmıştır. Nitel boyutta çalışma verilerini destekleyen öğretmen görüşlerinden metin içinde alıntı yapılarak araştırmanın

güvenirliliğine katkı sunulması hedeflenmiştir. Bu noktada yapılan alıntılara değinirken öğretmenlere 5Ö-e veya 7Ö-k (5 numaralı erkek öğretmen veya 2 numaralı kadın öğretmen) şeklinde kodlar verilerek araştırma boyunca kullanılmıştır.

BULGULAR

Bu bölümde çoklu analiz kapsamında ulaşılan nitel ve nicel veriler ayrı başlıklar altında sunulmuştur. sunmaktadır. Nicel bağlamda Rasch ve SPSS tabloları ile nitel bağlamda Maxqda modeli verilerek ilgili veriler yorumlanmıştır. Nitel ve nicel verilerin harmanlanarak birleştirilmesi ise sonuç kısmında genel sonuçlara ulaşma amacıyla yapılmıştır.

Rasch Modeli ile SPSS Verilerine İlişkin Bulgular

Rasch modeli yetenek düzeyleri, soruların güçlük durumu ve jüri katılımı/cömertliği şeklinde üç yüzeyden oluşmaktadır. Bu çok yüzeyli model, katılımcıların performans gerçekleştirme becerilerinin, soruların güçlük/kolaylık durumunun ve jüri katılım/cömertlik davranışlarının aynı anda karşılaştırılmasını sağlamaktadır (Akın ve Baştürk, 2012). Bu çalışmada İÖP'ün PÖDDEM'e göre değerlendirilmesine ilişkin görüşler çok-yüzeyli Rasch ölçme modeli ile analiz edilerek ilgili bilgiler Şekil 1'deki kalibrasyon haritasında sunulmuştur. Kalibrasyon haritasında "İÖP" (5., 6., 7. ve 8. sınıflar), "jüriler" ve "Kriterler" farklı sütunlarla gösterilmiştir. (-) ve (+) olarak gösterilen değerler, her yüzey için de aynı olan logit ölçüsüdür. İÖP'ün PÖDDEM bağlamındaki nitelikleri, jürilerin katılım/cömertlik puanlamaları ve maddelerdeki güçlük düzeyleri bu cetvelde logit ölçüsü ile sıralanmıştır. Programların nitelikleri bağlamındaki sıralamada, PÖDDEM'e göre en nitelikli program 7. Sn. İÖP 2 kodlu program iken niteliği en düşük program ise 5. Sn. İÖP 4 kodlu program olarak görülmektedir. Jüri puanlamalarına bakıldığında, en cömert jürinin J15 olduğu, J10 ve J26 kodlu jürilerin ise eşit düzeyde ve en katı jüri davranışı sergiledikleri anlaşılmaktadır. Jürilerin puanlamada birbirinden farklı davrandıkları görülmekte ancak eşit düzeyde davranış sergileyen jürilerin de bulunduğu dikkat çekmektedir. İÖP'ü PÖDDEM bağlamında değerlendirmek için kullanılan kriterler yüzeyine bakıldığında, öğretmenlerin gerçekleştirilmesi en zor madde olarak M1, M3 VE M9 kodlu maddeleri belirttikleri görülmektedir. M14, M15, M17, M20 VE M6 kodlu maddelerin ise gerçekleştirilmesi kolay olan maddeler olduğu anlaşılmaktadır.

Measr	+İng. Öğrt. Prog.	+Jürî	-Maddeler	S.1
+ 2 +		J2 J29		+(5) +
		J23 J30 J31 J35		4
		J1 J27 J34		
+ 1 +		J13 J26 J36 J37 J42 J7		+ +
		J12 J15 J21 J24 J25		
		J11 J5 J8		---
		J10 J14 J16 J20 J28 J39 J4		
		J18 J22 J33 J6		
		J17 J19 J38 J40		
		J3 J41		
	7. Sn. İÖP 2	J9	M1 M3 M9	3
	8. Sn. İÖP 1	J32	M2 M24	
			M10 M21	
* 0 *			M11 M13 M16 M18 M22 M5 M7 M8	* *
	6. Sn. İÖP 3		M12 M19 M23 M4	
	5. Sn. İÖP 4		M14 M15 M17 M20 M6	
+ -1 +				+(1) +
Measr	+İng. Öğrt. Prog.	+Jürî	-Maddeler	S.1

Sn: Sınıf

Şekil 1: Veri Kalibrasyon Haritası

Rasch analizlerinin yanı sıra faktörler öğretmenler tarafından belirtilen aritmetik ortalamalarının ne düzeyde olduğu SPSS Programıyla analiz edilerek daha detaylı bilgi sunulması amaçlanmıştır. Bu nedenle Rasch ölçme modeli ile SPSS programı sonuçlarının sınıflar bazında madde aritmetik ortalamaları karşılaştırıldığında, SPSS analizinde en yüksek ortalamanın 7. Sınıfta ($\bar{X}=3.86$), en düşük ise 5. Sınıfta ($\bar{X}=3.24$) olduğu saptanmıştır. Elde edilen bulgular, PÖDDEM açısından değerlendirildiğinde en nitelikli programın 7. sınıf, en düşük niteliğe sahip

programının ise 5. sınıf düzeyinde olduğu belirtilebilir. Dolayısıyla Rasch ve SPSS analizlerinin birbirini doğrulayıcı ve destekleyici bulgularıyla tutarlı sonuçlara ulaşıldığı anlaşılmıştır.

Tablo 1: 9-12. Sınıf İngilizce Dersi Öğretim Programlarına İlişkin Ölçüm Raporu

Obsvd Score	Obsvd Count	Obsvd Average	Fair Average	Model Measure	S.E.	Infit MnSq	Zstd	Outfit MnSq	Zstd	N İng. Öğrt. Prog.
3919	1008	3.9	3.91	.25	.03	1.0	1	1.1	1	2 7.Sn.İÖP 2
3882	1008	3.9	3.87	.21	.03	1.0	0	1.0	0	1 8.Sn.İÖP 1
3595	1008	3.6	3.58	-.06	.03	1.0	0	1.0	0	3 6.Sn.İÖP 3
3223	1008	3.2	3.19	-.40	.03	0.9	-2	0.9	-2	4 5.Sn.İÖP 4
3654.8	1008.0	3.6	3.64	.00	.03	1.0	-0.0	1.0	-0.1	Mean (Count: 4)
279.0	0.0	0.3	0.29	.26	.00	0.1	1.4	0.1	1.4	S.D.

RMSE (Model) .03 Adj S.D. .26 Separation 8.38 Reliability .99
Fixed (all same) chi-square: 288.2 d.f.: 3 significance: .00
Random (normal) chi-square: 3.0 d.f.: 2 significance: .22

Tablo 1’de İÖP’in (5-8. sınıflar) hedef, içerik, öğrenme durumları ve değerlendirme aşamalarına uygun nitelikleri kapsayan ayrıntılı bir ölçüm raporu verilmiştir. Programlar PÖDDEM aşamalarına göre en nitelikliden en az nitelikliye göre sıralanmıştır. Bu anlamda ilgili veriler en üstün nitelikli programın “7. Sn. İÖP 2” kodlu program; en düşük nitelikli olanın ise “5. Sn. İÖP 4” olduğunu göstermiştir. Bunun yanı sıra, Tablo 1’de programların nitelikleriyle ilgili logit değerlerine yönelik standart hata (RMSE) değerinin .03 olarak hesaplandığı görülmüştür. Bu değer aşırı uç değerlerin dışında bütün verilere yönelik ölçme hatasını göstermektedir. Değerin .03 olması, standart hatanın oldukça düşük olduğu şeklinde yorumlanabilir. Bu hata oranı dikkate alınarak düzeltilmiş standart sapma değeri de kritik değer olan 1.0’in altında bulunmuştur (Adj S.D.=.26). Rasch analizinde İÖP’in hangi güvenilirlikle PÖDDEM’e uygun nitelikte sıralandıkları güvenilirlik katsayısı ile belirlenmiştir. Rasch analizi ile .99 olarak belirlenen güvenilirlik katsayısı programların oldukça yüksek bir güvenilirlik düzeyiyle sıralandığını göstermektedir. Ayrıca tablodaki diğer veriler detaylı bir şekilde incelendiğinde, 8.38 Ayırma İndeksi, .99 güvenilirlik katsayısı ile kay-kare testi yapıldığında, İÖP’in hedef, içerik, öğrenme durumları ve değerlendirme aşamalarına uygunluk açısından istatistiksel olarak anlamlı farklılık gösterdiği belirtilebilir ($\chi^2=288.2$, $sd=3$, $p=0.00$).

Rasch analizinde ayrıca çalışmanın yüzeyleriyle ilgili “uygunluk içi (Infit)” ve “uygunluk dışı (Outfit)” istatistiksel değerler de sunulabilmektedir. Bu iki değer için kalite kontrol limiti 0.6 ile 1.4 aralığında belirlenmiştir (Wright ve Linacre, 1994). Uygunluk içi indeksi, karar sürecinde beklenmeyen cevaplar için bilgi, uygunluk dışı indeksi ise uzaktaki beklenmeyen cevaplar için (outlier) hassasiyet gösteren değerler (Baştürk, 2010) olarak ifade edilmiştir. Tablo 1’de, her iki indeksin belirlenen değerlerde olduğu görülerek sınırı aşan bir durum olmadığı anlaşılmıştır. Dolayısıyla İÖP’ü değerlendirmeye yönelik yapılan analizde belirlenen uygunluk içi ve uygunluk dışı değerlerin beklenen kalite kontrol değerleri içinde yer aldıkları söylenebilir.

Jüri analizi

Bu bölümde jürilerin İÖP’e ilişkin görüşlerinin katılık/cömertlik bilgileri sunulmuştur. Tablo 2’de program değerlendirmede en cömert davranışı sergileyen jürinin 404 puanla J2; en katının ise 288 puanla J32 olduğu görülmektedir. Diğer jürilerin davranış durumu Tablo 2’de en cömert puanlamadan en katı puanlamaya doğru sıralanıştan izlenebilmektedir. Diğer yandan jüri ayırma indeksinin 2.62 ve güvenilirlik katsayısının da 0.87 olduğu görülerek jüri puanlamalarının katılık/cömertlikleri arasında istatistiksel olarak anlamlı bir farklılık bulunduğu söylenebilir ($\chi^2=313.3$, $sd=41$, $p=0.00$).

Tablo 2: Jürilerin Katılık/Cömertlik Karşılaştırması

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Measure	Model S.E.	Infit Mnsq Zstd	Outfit Mnsq Zstd	Nu	Juri
404	96	4.2	4.24	1.43	.12	1.1	0	2	J2
402	96	4.2	4.22	1.40	.11	1.4	-2	29	J29
388	96	4.0	4.07	1.23	.11	0.4	-5	35	J35
385	96	4.0	4.04	1.19	.11	0.7	-2	30	J30
384	96	4.0	4.02	1.18	.11	0.7	-3	23	J23
384	96	4.0	4.02	1.18	.11	1.4	2	31	J31
380	96	4.0	3.98	1.13	.11	1.1	0	1	J1
376	96	3.9	3.94	1.09	.10	0.9	-1	27	J27
375	96	3.9	3.93	1.08	.10	0.3	-7	34	J34
371	96	3.9	3.88	1.04	.10	0.8	-1	42	J42
370	96	3.9	3.87	1.02	.10	0.9	-1	13	J13
364	96	3.8	3.81	.96	.10	0.7	-3	36	J36
364	96	3.8	3.81	.96	.10	0.6	-3	37	J37
363	96	3.8	3.80	.95	.10	1.4	2	7	J7
363	96	3.8	3.80	.95	.10	1.0	0	26	J26
359	96	3.7	3.75	.91	.10	1.2	1	21	J21
358	96	3.7	3.74	.90	.10	1.4	3	12	J12
355	96	3.7	3.71	.87	.10	1.2	1	15	J15
355	96	3.7	3.71	.87	.10	1.1	1	25	J25
353	96	3.7	3.69	.85	.10	1.2	1	24	J24
350	96	3.6	3.66	.82	.10	1.6	4	5	J5
344	96	3.6	3.59	.76	.10	1.0	0	8	J8
343	96	3.6	3.59	.75	.10	1.0	0	11	J11
337	96	3.5	3.53	.70	.10	0.7	-2	10	J10
337	96	3.5	3.53	.69	.10	1.1	0	4	J4
336	96	3.5	3.50	.68	.10	0.8	-1	28	J28
336	96	3.5	3.50	.68	.10	0.5	-5	14	J14
335	96	3.5	3.49	.67	.10	1.1	0	20	J20
335	96	3.5	3.49	.67	.10	1.0	0	16	J16
331	96	3.4	3.45	.64	.10	1.1	0	39	J39
329	96	3.4	3.43	.62	.10	1.0	0	6	J6
328	96	3.4	3.42	.61	.10	0.9	-1	22	J22
323	96	3.4	3.36	.56	.10	1.1	1	33	J33
318	96	3.3	3.31	.51	.10	0.9	0	18	J18
316	96	3.3	3.29	.49	.10	1.3	2	38	J38
315	96	3.3	3.27	.48	.10	1.1	0	40	J40
314	96	3.3	3.26	.47	.10	1.0	0	17	J17
309	96	3.2	3.21	.42	.10	1.0	0	19	J19
308	96	3.2	3.20	.41	.10	1.3	2	3	J3
296	96	3.1	3.07	.30	.10	1.4	3	41	J41
288	96	3.0	2.98	.22	.10	0.5	-5	9	J9
288	96	3.0	2.98	.22	.10	0.5	-5	32	J32

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Measure	Model S.E.	Infit Mnsq Zstd	Outfit Mnsq Zstd	Nu	Juri
348.1	96.0	3.6	3.63	.81	.10	1.0	-0.3	1.0	-0.3
27.9	0.0	0.3	0.30	.28	.00	0.3	2.7	0.3	2.6

RMSE (Model) .10 Adj S.D. .27 Separation 2.62 Reliability .87
Fixed (all same) chi-square: 313.3 d.f.: 41 significance: .00
Random (normal) chi-square: 40.7 d.f.: 40 significance: .44

Tablo 2’de ayrıca Rasch modeli yüzeylerinden jüri yüzeyinin “uygunluk içi” ve “uygunluk dışı” değerlerine bakıldığında, bazı jürilerin beklenen kalite kontrol değerleri (1.4 - 0.6 aralığı) arasında yer almadığı görülmektedir. Bu durum, jürilerin tutarlı davranmadıkları anlamına gelmektedir. J35, J34, J32 VE J20 kodlu jürilerin uygunluk içi ve uygunluk dışı kareler ortalamasının beklenen değerlerin dışında olduğu anlaşılmaktadır. Bu jürilerin beklenen puanlar yerine katı veya cömert yönde davrandıkları belirtilebilir.

Tablo 3: Cinsiyet Değişkenine Göre İngilizce Öğretmenlerinin Mevcut Programa İlişkin Görüşleri

Prog.	Faktörler	Erkek-1 (n=33;%42.3)		Bayan-2 (n=45;%57,7)		Levene Testi		t ve p Değerleri	
		\bar{X}	ss	\bar{X}	ss	F	p	t	p
8. Sın. İÖP (1.Ürün)	Hedef	3.66	.68	3.73	.62	.805	.373	-.520	.604
	İçerik	3.80	.65	3.85	.68	.196	.659	-.343	.732
	Öğrn.Dur.	3.78	.72	4.10	.64	3.129	.081	-2.072	.042
	Değ.	3.44	.69	3.75	.61	.635	.428	-2.097	.039
7. Sın. İÖP (2.Ürün)	Hedef	3.68	.58	3.95	.63	.218	.642	-2.062	.043
	İçerik	3.65	.48	3.92	.58	2.234	.139	-2.169	.033
	Öğrn.Dur.	3.81	.61	4.03	.44	3.675	.059	-1.831	.071
	Değ.	3.86	.60	3.87	.69	.219	.641	-.099	.921
6. Sın. İÖP (3.Ürün)	Hedef	3.28	.41	3.47	.39	1.089	.300	-2.116	.038
	İçerik	3.33	.51	3.60	.57	.004	.948	-2.096	.039
	Öğrn.Dur.	3.60	.61	3.75	.66	.328	.569	-1.027	.308
	Değ.	3.64	.75	3.79	.73	.173	.678	-.873	.385
5. Sın. İÖP (4.Ürün)	Hedef	3.13	.59	3.11	.45	1.139	.289	-.213	.832
	İçerik	3.00	.58	3.32	.67	.922	.340	-2.207	.030
	Öğrn.Dur.	3.12	.64	3.43	.68	.084	.772	-2.028	.046
	Değ.	3.13	.66	3.52	.64	.568	.453	-2.744	.008

P<.05 df:76

Diğer yandan araştırmada Rasch analizinin SPSS programı çıktılarıyla karşılaştırılması sürecinde, ölçme aracına ilişkin katılımcıların cinsiyet değişkenine göre düşüncelerinde, sınıf bazında anlamlı farklılık olup olmadığı SPSS programı ile analiz edilmiştir. Katılımcıların sınıf bazında cinsiyet değişkeni açısından değerlendirilmesine yönelik veriler Tablo 3'te sunulmuştur. İÖP'ün t-testi analizi sonucunda içerik faktöründe 8. sınıf düzeyi hariç diğer tüm sınıflarda cinsiyet değişkeni açısından İngilizce öğretmenlerinin görüşleri arasında anlamlı fark ortaya çıkmıştır ($p < .05$). Tabloda anlamlı farklılığın kadın öğretmenlerin ($\bar{X}_{7.sın}=3.92$, $\bar{X}_{6.sın}=3.60$, $\bar{X}_{5.sın}=3.32$) lehine olduğu anlaşılmıştır. Bu bulgu, kadın öğretmenlerin içerik faktörüne yönelik katılım düzeylerinin diğer sınıf düzeylerinden daha yüksek olduğunu göstererek Rasch analizi sonuçlarıyla aynı yönde sonuçlara ulaşmıştır. Tablo 3'e bakıldığında, genel anlamda cinsiyet değişkeninin toplam sınıflar bazında değerlendirilmesi yapıldığında, yine kadın öğretmenlerin ($\bar{X}_{8.sın}=3.86$; $\bar{X}_{7.sın}=3.95$; $\bar{X}_{6.sın}=3.65$; $\bar{X}_{5.sın}=3.34$) tüm sınıf düzeylerinde erkek öğretmenlere ($\bar{X}_{8.sın}=3.67$; $\bar{X}_{7.sın}=3.75$; $\bar{X}_{6.sın}=3.47$; $\bar{X}_{5.sın}=3.09$) oranla daha yüksek düzeyde katılım gösterdikleri anlaşılmıştır. Ayrıca bu değerlere bakıldığında en yüksek değerleri alan sınıf düzeyinin 7. Sınıf İÖP olduğu görülmüştür.

İÖP'in PÖDDEM'e Göre Değerlendirmesinde Kullanılan Ölçek Maddelerinin Analizi

Araştırmada İngilizce öğretmenlerinin İÖP'i PÖDDEM bağlamında değerlendirmeleri amacıyla kullanılan değerlendirme formu maddelerinin amaca ne derece uygun olup olmadığı incelenerek Tablo 4'te sunulmuştur. Tabloda madde güçlük analizine ilişkin ayırma indeksinin 1.75 ve güvenilirlik katsayısının .75 olduğu görülmektedir. Güvenirlik katsayısı maddelerin değerlendirme yapmak için güvenilir olduğunu göstermiştir. Diğer yandan madde güçlüklerine bakıldığında, en zor gerçekleştirilen maddenin M3; en kolay gerçekleştirilen maddenin ise M17 kodlu maddeler olduğu görülmüştür. Araştırmanın yüzeyleri ile ilgili uygunluk içi ve uygunluk dışı istatistiksel değerlere bakıldığında, belirlenen sınır değerlerini (1.6-0.4) aşan değerlere rastlanmamıştır. Bu bulgu program değerlendirmesi için kullanılan maddelerin tutarlı oldukları anlamına gelmektedir.

Tablo 4: Ortaöğretim İngilizce Programların Değerlendirilmesinde Kullanılan Yeterlik İstatistikleri

Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Measure	Model S.E.	Infit Mnsq	ZStd	Outfit Mnsq	ZStd	Nu Maddeler
553	168	3.3	3.28	.32	.07	1.2	1	1.2	1	3 M3
560	168	3.3	3.33	.28	.07	1.0	0	1.0	0	9 M9
563	168	3.4	3.35	.26	.07	1.1	1	1.1	1	1 M1
573	168	3.4	3.41	.21	.07	1.1	0	1.1	0	2 M2
582	168	3.5	3.47	.16	.08	0.8	-2	0.8	-2	24 M24
594	168	3.5	3.55	.09	.08	1.0	0	1.0	0	10 M10
599	168	3.6	3.58	.06	.08	1.2	2	1.2	2	21 M21
601	168	3.6	3.59	.05	.08	0.9	-1	0.8	-1	7 M7
604	168	3.6	3.61	.03	.08	1.0	0	1.0	0	8 M8
605	168	3.6	3.62	.03	.08	0.8	-2	0.8	-2	18 M18
608	168	3.6	3.64	.01	.08	0.8	-2	0.8	-2	13 M13
611	168	3.6	3.66	-.01	.08	1.1	0	1.1	0	5 M5
613	168	3.6	3.67	-.02	.08	1.0	0	1.0	0	22 M22
614	168	3.7	3.67	-.03	.08	0.8	-2	0.8	-2	16 M16
618	168	3.7	3.70	-.05	.08	1.2	1	1.2	1	11 M11
625	168	3.7	3.74	-.09	.08	1.1	1	1.2	1	4 M4
626	168	3.7	3.75	-.10	.08	1.0	0	1.0	0	23 M23
630	168	3.8	3.78	-.12	.08	0.9	0	0.9	0	12 M12
632	168	3.8	3.79	-.13	.08	0.8	-1	0.9	-1	19 M19
636	168	3.8	3.81	-.16	.08	1.0	0	1.0	0	15 M15
641	168	3.8	3.85	-.19	.08	1.3	3	1.3	3	6 M6
641	168	3.8	3.85	-.19	.08	0.9	-1	0.9	-1	14 M14
644	168	3.8	3.86	-.20	.08	1.0	0	1.0	0	20 M20
646	168	3.8	3.88	-.22	.08	1.1	0	1.1	0	17 M17
Obsvd Score	Obsvd Count	Obsvd Average	Fair Avrage	Measure	Model S.E.	Infit Mnsq	ZStd	Outfit Mnsq	ZStd	Nu Maddeler
609.1	168.0	3.6	3.64	.00	.08	1.0	-0.1	1.0	-0.1	Mean (Count: 24)
26.7	0.0	0.2	0.17	.15	.00	0.1	1.6	0.1	1.6	S.D.

RMSE (Model) .08 Adj S.D. .13 Separation 1.75 Reliability .75
Fixed (all same) chi-square: 98.1 d.f.: 23 significance: .00
Random (normal) chi-square: 23.0 d.f.: 22 significance: .40

Jüri Yanlılığı Etkileşim Analizi

Tablo 5'te İÖP'ün PÖDDEM'e yönelik jüriler tarafından yapılan görüşlerin etkileşim analizi verilmiştir. Bu veriler incelendiğinde, J5 kodlu jürinin 8. Sn. İÖP kodlu programa 93.1 puan vermesi gerekirken 70 puan verdiği ve katı yönde yanlı bir davranış sergilediği ($Z=4.41$); J38'in ise aynı programa 85.2 puan vermesi beklenirken 111 puan verdiği ve cömert yönde yanlı bir davranış sergilediği söylenebilir ($Z=-4.31$). Bununla birlikte, Z puanlarının +2 ve

-2 dışında farklı değerler alması bir etkileşim yanlılığı anlamına gelmektedir (Semerci, 2011). Mevcut araştırmada Z puanlarının -4.31 ile 4.41 değerleri arasında olduğu görülmektedir. Tablo 5'te yanlılık analizi açısından veriler incelendiğinde, jürilerin İÖP'ü değerlendirmede aşırı derecede katı ya da cömert davrandıkları belirtilebilir. Dolayısıyla farklı sebeplere dayandırılabilir, cömert ve katı davranış sergileyen jürilerin yanlı davrandıkları söylenebilir.

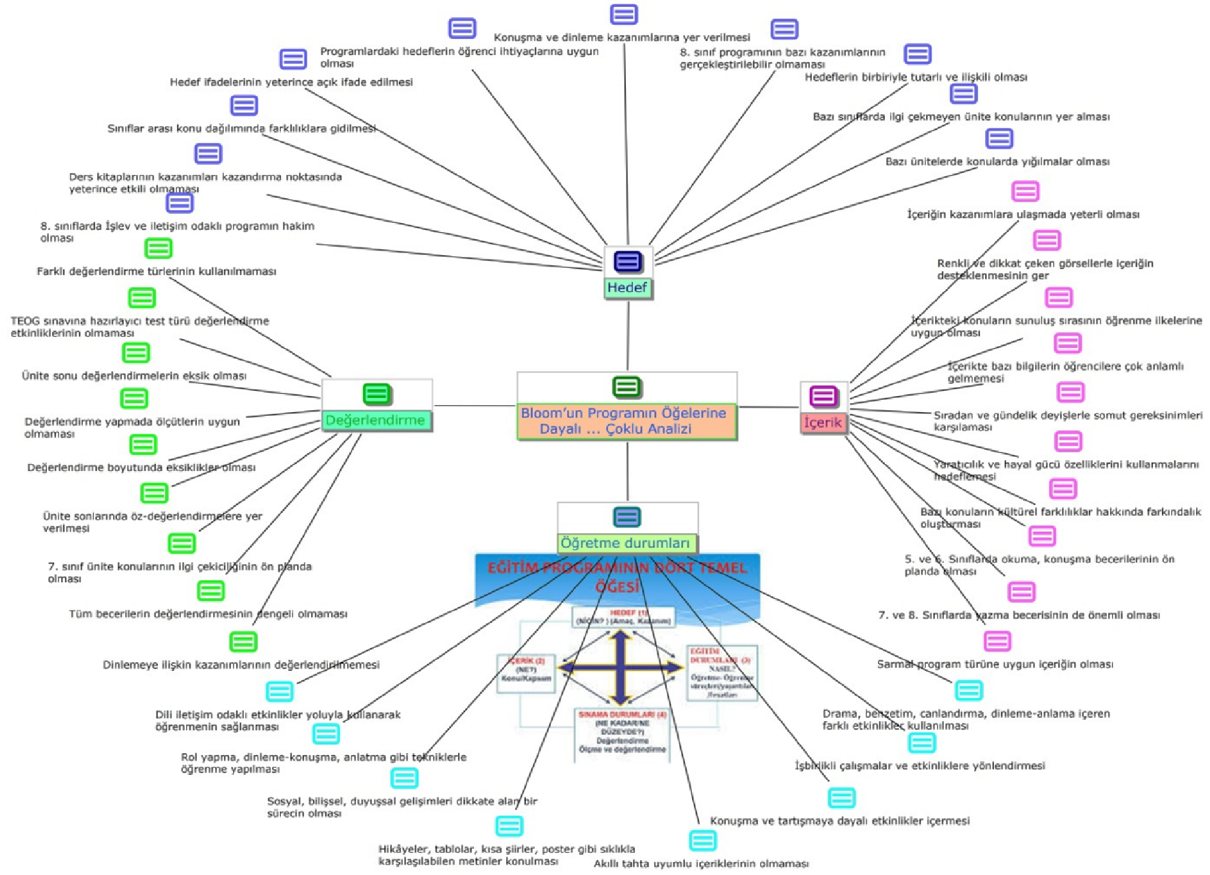
Tablo 5: Jüriler ile İÖP'ün Etkileşim Analizi

Obsvd Score	Exp. Score	Obsvd Count	Obs-Exp Average	Bias+ Measure	Model S.E.	Z-Score	Infit MnSq	Outfit MnSq	Sq	N	İng.Öğ.Prog.measr	Nu	Puan	measr	
70	93.1	24	-.96	.88	.20	4.41	1.6	1.6	17	1	8.Sn.İÖP 1	.21	5	J5	.82
60	82.6	24	-.94	.93	.22	4.13	0.4	0.4	15	3	6.Sn.İÖP 3	-.06	4	J4	.69
62	82.9	24	-.87	.84	.22	3.84	0.5	0.5	9	1	8.Sn.İÖP 1	.21	3	J3	.42
68	86.2	24	-.76	.69	.20	3.40	0.7	0.7	150	2	7.Sn.İÖP 2	.25	38	J38	.51
56	73.5	24	-.73	.80	.24	3.34	0.7	0.7	16	4	5.Sn.İÖP 4	-.40	4	J4	.69
73	90.4	24	-.73	.66	.20	3.34	1.3	1.3	62	2	7.Sn.İÖP 2	.25	16	J16	.67
60	77.6	24	-.73	.74	.22	3.30	1.4	1.5	96	4	5.Sn.İÖP 4	-.40	24	J24	.85
67	84.7	24	-.74	.68	.21	3.30	1.7	1.7	157	1	8.Sn.İÖP 1	.21	40	J40	.49
52	68.0	24	-.67	.82	.26	3.21	0.3	0.3	76	4	5.Sn.İÖP 4	-.40	19	J19	.47
70	86.4	24	-.69	.62	.20	3.10	1.2	1.2	69	1	8.Sn.İÖP 1	.21	18	J18	.56
87	100.7	24	-.57	.59	.20	3.03	1.0	1.1	2	2	7.Sn.İÖP 2	.25	1	J1	1.13
89	101.8	24	-.53	.57	.20	2.90	0.5	0.5	118	2	7.Sn.İÖP 2	.25	30	J30	1.19
63	78.2	24	-.63	.62	.22	2.86	0.6	0.6	100	4	5.Sn.İÖP 4	-.40	25	J25	.87
82	95.2	24	-.55	.52	.19	2.69	1.1	1.1	81	1	8.Sn.İÖP 1	.21	21	J21	.91
70	84.2	24	-.59	.54	.20	2.68	0.8	0.8	43	3	6.Sn.İÖP 3	-.06	11	J11	.75
95	105.4	24	-.43	.54	.21	2.58	1.1	1.1	114	2	7.Sn.İÖP 2	.25	29	J29	1.40
82	94.3	24	-.51	.48	.19	2.48	1.3	1.3	57	1	8.Sn.İÖP 1	.21	15	J15	.87
96	105.8	24	-.41	.52	.21	2.46	0.6	0.6	6	2	7.Sn.İÖP 2	.25	2	J2	1.43
60	72.1	24	-.51	.53	.22	2.37	1.3	1.4	35	3	6.Sn.İÖP 3	-.06	9	J9	.30
61	73.2	24	-.51	.52	.22	2.37	0.9	0.9	56	4	5.Sn.İÖP 4	-.40	14	J14	.68
81	91.6	24	-.44	.40	.19	2.09	0.7	0.7	167	3	6.Sn.İÖP 3	-.06	42	J42	1.04
92	101.1	24	-.38	.41	.20	2.06	0.4	0.4	117	1	8.Sn.İÖP 1	.21	30	J30	1.19
86	95.9	24	-.41	.40	.19	2.04	1.5	1.5	46	2	7.Sn.İÖP 2	.25	12	J12	.90
77	66.8	24	-.42	-.40	.19	-2.08	0.5	0.5	12	4	5.Sn.İÖP 4	-.40	3	J3	.42
105	94.7	24	.43	-.53	.25	-2.09	0.7	0.6	94	2	7.Sn.İÖP 2	.25	24	J24	.85
99	88.0	24	.46	-.47	.22	-2.12	1.0	1.0	85	1	8.Sn.İÖP 1	.21	22	J22	.62
110	99.8	24	.43	-.67	.30	-2.21	1.4	1.4	115	3	6.Sn.İÖP 3	-.06	29	J29	1.40
96	84.1	24	.49	-.47	.21	-2.24	0.9	1.0	73	1	8.Sn.İÖP 1	.21	19	J19	.47
99	86.8	24	.51	-.51	.22	-2.33	0.4	0.4	120	4	5.Sn.İÖP 4	-.40	30	J30	1.19
88	75.4	24	.52	-.47	.20	-2.40	0.7	0.7	11	3	6.Sn.İÖP 3	-.06	3	J3	.42
90	76.8	24	.55	-.49	.20	-2.50	1.1	1.1	20	4	5.Sn.İÖP 4	-.40	5	J5	.82
112	100.0	24	.50	-.87	.34	-2.55	0.6	0.6	1	1	8.Sn.İÖP 1	.21	1	J1	1.13
108	94.3	24	.57	-.75	.28	-2.70	1.5	1.5	97	1	8.Sn.İÖP 1	.21	25	J25	.87
105	90.6	24	.60	-.69	.25	-2.75	0.9	1.0	54	2	7.Sn.İÖP 2	.25	14	J14	.68
107	92.6	24	.60	-.75	.27	-2.79	1.1	1.1	30	2	7.Sn.İÖP 2	.25	8	J8	.76
109	92.3	24	.69	-.92	.29	-3.15	1.1	1.2	42	2	7.Sn.İÖP 2	.25	11	J11	.75
110	90.9	24	.80	-1.06	.30	-3.49	1.1	1.1	14	2	7.Sn.İÖP 2	.25	4	J4	.69
111	89.9	24	.88	-1.20	.32	-3.74	1.0	0.9	13	1	8.Sn.İÖP 1	.21	4	J4	.69
101	79.4	24	.90	-.89	.23	-3.88	1.1	1.2	33	1	8.Sn.İÖP 1	.21	9	J9	.30
111	85.2	24	1.08	-1.38	.32	-4.31	0.6	0.7	149	1	8.Sn.İÖP 1	.21	38	J38	.51
Obsvd Score	Exp. Score	Obsvd Count	Obs-Exp Average	Bias+ Measure	Model S.E.	Z-Score	Infit MnSq	Outfit MnSq	Sq	N	İng.Öğ.Prog.measr	Nu	Puan	measr	
87.0	87.0	24.0	.00	-.02	.21	.01	0.9	0.9	Mean (Count: 168)						
13.2	9.6	0.0	.37	.40	.03	1.73	0.9	0.4	S.D.						

Fixed (all = 0) chi-square: 503.9 d.f.: 168 significance: .00

Araştırmanın Nitel Verilerine İlişkin Bulgular

Çalışmanın bu bölümünde katılımcılara görüşme formuyla yöneltilen program değerlendirmeyle ilgili sorulara ilişkin görüşler sunulmuştur. Maxqda programıyla çözümlenen veriler model haline getirilerek aşağıda Şekil 2'de verilmiştir. Model incelendiğinde araştırmanın nitel boyutuna ilişkin bulgular, hedef, içerik, öğrenme durumları ve değerlendirme şeklinde dört tema altında toplanmıştır.



Şekil 2: Programın Öğelerine Dayalı Değerlendirme Modeli Bağlamında İÖP'e İlişkin Görüşler

Şekil 2'de İÖP'ün hedef teması bağlamında dile getirilen kodlardan bazıları "konuşma ve dinleme kazanımlarına yer vermesi, bazı sınıflarda ilgi çekmeyen ünite konularının olması, hedef ifadelerinin yeterince açık olması" şeklindedir. Bu kodlara referans olabilecek ifadeler 3Ö-k tarafından "öğretim programı tüm sınıflar bazında incelendiğinde aslında yeterince dinleme ve konuşma etkinliklerine yer verildiği, kazanımlar da bu becerilerin de i-kazandırılması gerektiği vurgulanmaktadır" biçiminde ve 8Ö-e tarafından "Programlar elbette tüm sınıf düzeyleri için aynı değil. Örneğin 8. Sınıf 3. Ünite konusunun "yemek pişirme" ile ilgili olması özellikle erkek öğrencilerin hiç ilgisini çekmeyerek dersten sıkılmalarına neden olabiliyor." biçiminde belirtilmiştir. Diğer yandan içerik bağlamında dile getirilen kodlardan bazıları "içeriğin kazanımlara ulaşmada yeterli olması, yaratıcılık ve hayal gücünü kullanmalarını sağlaması, Renkli ve dikkat çeken görsellerle içeriğin desteklenmesinin gerekmesi" olarak görülmektedir. Bu kodlar 13Ö-k kodlu öğretmenin "...genel anlamda ders içeriğinin yeterli olduğu kanısındayım. Farklı becerilere yönelik belirlenen metin, diyalog, şarkı gibi içeriklerin amaçlanan kazanımlara ulaştırmada iyi olduğunu düşünüyorum" şeklindeki ifadesi; 8Ö-e kodlu öğretmenin ise "Programın daha çok ilgi çeken içeriklerle donatılması gerekir. Öğrencilerin ilköğretim düzeyinde olmaları, onların daha renkli, ilginç ve dikkat çeken materyallere ilgi duymaları anlamına gelmektedir. Bu nedenle içeriğin daha zengin hazırlanması gerekir." biçimindeki cümlesi esas alınarak oluşturulmuştur.

Programın öğretme durumları bağlamında değerlendirilmesi sonrasında "Rol yapma, dinleme-konuşma, anlatma gibi tekniklerle öğrenme yapılması, Hikâyeler, tablolar, kısa şiirler, poster gibi sıklıkla karşılaşılabilen metinler konulması" şeklindeki kodlara ulaşıldığı görülmektedir. Bu kodların alıntılı olduğu ifadelerden biri 15Ö-k kodlu katılımcının "öğretim programının farklı teknik ve stratejiler içermesi öğrenme sürecini verimli kılmaktadır. Bazen hikâye veya şiir okuma, bazen poster hakkında konuşma, bazen de mektup yazma, farklı yollar kullanarak etkililiği arttırmaktadır." biçimindeki görüşü olarak belirtilebilir. Değerlendirme boyutunda ise "TEOG sınavına hazırlayıcı test türü değerlendirme etkinliklerinin olmaması, 7. sınıf ünite konularının ilgi çekiciliğinin ön planda olması, Ünite sonlarında öz-değerlendirmelere yer verilmesi" şeklindeki kodların olduğu görülmektedir. İlgili kodlara referans olabilecek görüşlerden biri 22Ö-k kodlu katılımcının "7. Sınıf konularının öğrencilere göre olduğunu, ünite sonlarındaki öz-değerlendirmelerin öğrencileri bilinçlendirdiğini ve

değerlendirme noktasında genel anlamda tüm sınıf düzeylerinde eksiklikler olduğunu söyleyebilirim.” cümlesinden alıntılanmıştır.

SONUÇ VE TARTIŞMA

Bu araştırma farklı analiz programları kullanılarak analizlerden birinde olan eksikliğin diğeri ile tamamlanmasını sağlamayı içeren çoklu analiz yaklaşımı doğrultusunda yürütülmüştür. Çoklu analiz ile ulaşılması amaçlanan farklı bulgulara uygun analiz programları ve veri kaynakları kullanılarak ulaşılabilir. Bu noktada veri kaynaklarının zenginliğinin farklı bakış açıları kazandırarak araştırmanın niteliğine katkı sağladığı bilinmektedir. Dolayısıyla hem nitel hem de nicel boyutlarda yürütülen araştırmada nicel anlamda Rasch ve SPSS programları ve nitel anlamda ise tematik analize dayalı Maxqda programında yararlanılmıştır. Rasch modeli sonucunda ulaşılan bulgulara göre 7. Sn. İÖP kodlu programın en nitelikli, 5. Sn. İÖP kodlu programın ise en düşük nitelikli olduğu anlaşılmıştır. SPSS analizleri sonucunda elde edilen bulgulara bakıldığında, sınıflar bazında katılımcıların yaptıkları değerlendirmede genel olarak en yüksek katılımın 7. Sınıf İÖP'e yönelik olduğu görülmüştür. Bu sonuç SPSS analizinde en nitelikli programın 7. Sınıf programı olduğunu göstermiştir. Dolayısıyla ulaşılan bu sonucun Rasch analizi çıktıları ile örtüştüğü, iki sonucun tutarlı olduğu ve birbirini tamamladığı söylenebilir. Jürilerin puanlamalarına ilişkin değerlendirmelere bakıldığında, J2 kodlu jürinin en cömert, J32 kodlu jürinin ise en katı jüri davranışına sahip olduğu söylenebilir. Dolayısıyla jürilerin puanlamada yanlı davrandıkları ve beklenen puanların dışında puanlamada buldukları söylenebilir. Uygunluk içi ve uygunluk dışı istatistiksel değerlerde tüm jürilerin, beklenen kalite kontrol değerleri (1.4-0.6 aralığı) arasında yer almadıklarının görülmesi yanlı davrandıklarını ve tutarlı puanlayıcı davranışı sergilemediklerini göstermektedir. Rasch analizinde yanlı davranış sergilenmesi durumu sadece puanların sıraya koyulması ile en cömert ve en katı şeklinde ifade edilmiştir. Ancak bu yanlılığın cinsiyet değişkenine göre incelenmesi Rasch analizinde olmadığından dolayı, SPSS analizi yapılması düşünülmüştür. Yapılan SPSS analizi sonucunda kadın ve erkek öğretmenler arasında anlamlı farklılıklara rastlanmıştır ve bu farkın kadın öğretmenler lehine olduğu anlaşılmıştır.

Araştırmada İÖP'ün PÖDDEM bağlamında değerlendirilmesine ilişkin hazırlanan formdaki madde güçlük analizi, güvenilirlik katsayısının .75 olduğunu göstermiştir. Bu değer, maddelerin program değerlendirmede güvenilir nitelikte olduğu şeklinde yorumlanabilir. Diğer yandan madde güçlük düzeylerine bakıldığında, en zor gerçekleştirilen maddenin M3 ve en kolay gerçekleştirilen maddenin ise M17 kodlu maddeler olduğu anlaşılmaktadır. Ayrıca maddelerin uygunluk içi ve uygunluk dışı değerlerinin belirlenen sınır aralıklarında olduğu saptanmıştır. Bu bulgu değerlendirme formunda kullanılan maddelerin tutarlı olduğu anlamına gelmektedir. Araştırmanın nicel boyutunda Rasch analiziyle ulaşılan verilerden sonuncusu ise çalışmanın güvenilirliğiyle ilgilidir. Bu bağlamda İÖP'ün niteliği noktasında .99; puanlayıcı düzeyi noktasında .87 ve madde kalitesi noktasında .75 değerinde güvenilirlik katsayılarına ulaşılmıştır. Bu değerlerin +1.00 yakınlarında olması güvenliliğin artması şeklinde yorumlanmıştır (Baştürk, 2010). Dolayısıyla çalışmanın üç yüze açısından da güvenilir düzeyde olduğu söylenebilir.

Araştırmaya veri zenginliği sağlamak amacıyla eklenen Maxqda verileri, Rasch ve SPSS analizinde ortaya çıkan verileri genel anlamda desteklemiştir. Bu noktada katılımcılardan elde edilen bulguların hedef, içerik öğrenme durumları ve değerlendirme temaları altında farklı kodlardan oluştuğu görülmüştür. Hedef teması bağlamında hedeflerin yeterinde açık ifade edildiği ve dinleme ve konuşma becerilerine de kısmen yer verildiği belirtilmiştir. Bu noktada Topkaya ve Küçük'ün (2010) 4. ve 5. Sınıf İngilizce programını değerlendirmeye yönelik çalışmalarında, dört temel dil becerilerinin programda dengeli dağılmadığı ve kazanımların yeterli netlikte ifade edildiği sonucu mevcut araştırmayı desteklemektedir. Diğer yandan içerik bağlamında, programdaki içeriğin kazanımlara ulaşmada yeterli olduğu, yaratıcılığı desteklediği ancak mevcut içeriğin daha zenginleştirilmesi ve çeşitlendirilmesi gerektiği belirtilmiştir. Öğrenme durumları kapsamında ise, programların çeşitli strateji ve tekniklerle ders işlenmesine imkân verecek yapıda düzenlendiği anlaşılmıştır. Özellikle farklı türdeki içeriklerin, işbirlikli ve grupsal çalışmaların bu noktada programların etkili yönleri olduğu vurgulanmıştır. Son olarak programların değerlendirme açısından etkililiğine ilişkin görüşlere başvurulduğunda, bu konuda programların yetersiz olduğu belirtilmiştir. Özellikle 8. Sınıflarda TEOG sınavına hazırlık noktasında ekstra hiçbir değerlendirme etkinliğinin olmaması büyük bir eksikli olarak görülmüştür. Diğer yandan öz-değerlendirme bölümlerine yer verilmesinin öğrencilerin kendilerini tanıma ve kendilerinin farkına varma noktasında olumlu etki ettiği söylenebilir. Araştırmanın nitel boyutu bağlamında programların genelde hedef, içerik ve öğrenme durumları açısından yeterli ancak değerlendirme faktörü açısından zayıf nitelikte olduğu belirtilmiştir. Bu

sebeple programların özellikle değerlendirme noktasında geliştirilmesi ve yeniden gözden geçirilmesi gerektiği önerilebilir.

Çoklu analiz sonucunda ulaşılan sonuçlar 7. Sınıfın nitelikli; ancak 5. Sınıfın düşük nitelikli program olduğunu göstermiştir. Dolayısıyla ilgili eksiklik ve yetersizliklerin tespit edilerek bu noktada gerekli iyileştirmelerin yapılması tavsiye edilmektedir. Program hazırlayan uzman ve tasarımcıların öğrencilerin yaş, cinsiyet, yaşadıkları çevre gibi özellikleri düşünerek hedefleri belirlemeleri, diğer program öğelerini tamamen hedefleri dikkate alarak planlamaları gerekir. Çeliköz (2004) bu noktada hedef belirlemenin öğrenciye göre olması gerektiğini; masa başında değil ilgililerin görüşleri doğrultusunda hazırlanmasını ifade etmiştir. Sonuç olarak çoklu analiz ile yürütülen araştırmada Rasch, SPSS ve Maxqda analiz verilerinin birbiriyle tutarlı ve birbirini tamamlayıcı sonuçlara ulaştığı görülmüştür. Özellikle Rasch ve SPSS sonuçlarında en nitelikli program olarak 7. Sınıf, en düşük nitelikli ise 5. sınıf programının saptanması iki araştırma sonucunun birbirini desteklediği ve tamamladığı anlamına gelmektedir. Diğer yandan nitel bulgularda ulaşılan ve 7. Sınıf programına yönelik daha fazla olumlu görüşün olması nitel bulguların da nicel bulgularla örtüştüğü şeklinde yorumlanabilir. Çoklu analizle ulaşılan bulguların birbiriyle aynı yönde olması ve tutarlı sonuçlar oluşturması araştırmanın güvenilirliğine ve niteliğine katkı sunmuştur. Bu tür çok boyutlu araştırmaların gelecekteki diğer araştırmalara farklı boyutlar kazandırması düşünülerek çoklu analizle farklı çalışmalar yapılması önerilmektedir.

Not: Bu çalışma 27- 29 Ekim 2016 tarihlerinde Antalya'da 7 Ülkenin katılımıyla düzenlenen World Conference on Educational and Instructional Studies- WCEIS'de bildiri olarak sunulmuştur.

KAYNAKÇA

Adamson, K. A. & Prion, S. (2015). Making sense of methods and measurement: simulation program evaluation. *Clinical Simulation in Nursing*, 11(1), 505-506.

Akın, Ö. & Baştürk, R. (2012). Keman eğitiminde temel becerilerin Rasch Ölçme Modeli ile değerlendirilmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 31, 175-187.

Andersn, L. W. (2005). Objectives, evaluation, and the improvement of education. *Studies in Educational Evaluation* 31, 102-113.

Bamberger, M. (2012). Introduction to mixed methods in impact evaluation. *Impact Evaluation Notes*, 3, 1-38.
Baştürk, R. (2010). Bilimsel araştırma ödevlerinin çok yüzeysel Rasch ölçme modeli ile değerlendirilmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(1), 51-57.

Boone, W.; Scantlebury; K. (2006). The role of Rash analysis when conducting science education research utilizing multiple-choice tests. *Science Education*, 90, 253-269.

Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2). pp. 77-101. ISSN:1478-0887. Retrieved from: <http://eprints.uwe.ac.uk/11735>.

Bümen, N. T. (2006). Program geliştirmede bir dönüm noktası: yenilenmiş Bloom taksonomisi. *Eğitim ve Bilim*, 31(142), 3-14.

Çeliköz, N. (2004). Yeni program geliştirme anlayışına dayalı olarak geliştirilen bir program tasarımının öğrenci başarısına etkisi. *GÜ, Gazi Eğitim Fakültesi Dergisi*, 24(1), 99-113.

Demirel, Ö (1992). Türkiye'de program geliştirme uygulamaları. *H. Ü. Eğitim Fakültesi Dergisi*, 7, 27-43.

Ertürk, S. (1998). *Eğitimde program geliştirme*. Ankara: Meteksan A.Ş.

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). *Program evaluation: Alternative approaches and practical guidelines*. NY: Pearson Education.

- Gökmenoğlu, T. (2014). Geniş Açı: Modeller ve Yaklaşımlar Açısından Türkiye’de Program Değerlendirme Çalışmaları. *Uluslararası Eğitim Programları ve Öğretim Çalışmaları Dergisi*, 4(7), 56-70.
- Gredler, M. E. (1996). *Program evaluation*. NJ: Pearson Education.
- Hubbard, W. & Bauer, D. (1993). Programme evaluation: What we can learn from the United States experience. *Australian Physiotherapy*, 39(1), 53-57.
- Kim, H. S. (2012). Outcomes-based curriculum development and student evaluation in nursing education. *J Korean Acad Nurs.*, 42(7), 917-927.
- Marvasti, A. B. (2004). *Qualitative research in sociology*. London: Sage Publications Ltd.
- Oliva, P. F. (1997). *Developing the curriculum*. New York: Longman.
- Özdamar, K. (2010). Paket programlar ile istatistiksel veri analizi 2. (7. edition). Eskişehir, Kaan Yayıncılık.
- Özdemir, S. M. (2009). Eğitimde program değerlendirme ve Türkiye’de eğitim programlarını değerlendirme çalışmalarının incelenmesi. *Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi Dergisi*, 6(2), 126-149.
- Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). London: Sage Publications, Inc.
- Pekdağ, B. ve Erol, H. (2013). 1957-2007 yılları arasında yayımlanan ortaöğretim Kimya dersi öğretim programlarının gerekçe, amaç ve içerik yönünden incelenmesi. *Kuram ve Uygulamada Eğitim Bilimleri*, 13(1), 631-659.
- Semerci, Ç. (2011): Mikro öğretim uygulamalarının çok yüzeysel Rasch ölçme modeli ile analizi. *Eğitim ve Bilim*, 36(161), 14-25.
- Stassen, M. L. A., Doherty, K. ve Poe, M. (2001). Program-based review and assessment. Campus publications designed by the Office of Academic Planning and Assessment (OAPA). UMass Amherst.
- Stufflebeam, D. L., & Shinkfield, A. J. (1985). *Systematic evaluation*. Boston: Kluwer-Nijhoff.
- Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of practical program evaluation*. San Francisco: Jossey-Bass.
- Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practice guidelines* (2nd ed.). New York: Longman.
- Yıldırım, A. ve Şimşek, H. (2006). *Sosyal bilimlerde nitel araştırma yöntemleri*, (5. Baskı), Ankara: Seçkin Yayıncılık.
- Yin, R., (1994). *Case study research: Design and methods* (2nd ed.). Beverly Hills, CA: Sage Publishing.
- Yüksel, I. (2010). *Türkiye için program değerlendirme standartları oluşturma çalışması*. Yayımlanmamış doktora tezi, Anadolu Üniversitesi, Eskişehir.
- Zainal, Z. (2007). Case study as a research method. *Jurnal Kemanusiaan*, 9, 1-6.
- Zehir Topkaya, E. ve Küçük, Ö. (2010). An evaluation of 4th and 5th grade English language teaching Program. *Elementary Education Online*, 9(1), 52-65.

Zikmund, W. G. (2003). Research methods. health economics research method, 2. Retrieved August 28, 2016 from <http://pioneer.netserv.chula.ac.th/~ppongsa/2900600/LMRM02.pdf>

Wright B. D. & Linacre J. M. (1994). Reasonable mean-square fit values. Transactions of the Rasch measurement SIG American Educational Research Association. *Rasch Measurement Transaction* 8(3), 370-382.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement:The kappa statistic. *Family Medicine*, 37(5), 360-363.

EK-1: Çoklu Analiz (Rasch, SPSS ve Maxqda) Gerekçeleri

Aşamalar	Rasch Ölçme Modeli	SPSS ve Rasch Programlarının Karşılaştırılması	Programların Birlikte Kullanılmasının Gerekçeleri
1.	Kalibrasyon haritası (Tüm yüzeylerin karşılaştırmalı analizi)	Üç yüzeyin özeti SPSS ve Rasch sonuçları şeklinde bu bölümde karşılaştırılmıştır.	SPSS Programı tek ürün değerlendirme odaklıdır. Rasch'ta pek çok ürün vardır. Katılımcıların tüm bunlar hakkındaki görüşlerinin bütüncül bir şekilde görülmesi istenmiştir. Yani SPSS'te katılımcıların x hakkındaki görüşleri analiz edilirken, Rasch'ta x, y, z gibi ürünler hakkındaki görüşleri incelenmektedir. Bu aşamada Rasch'ta bütüncül olarak tüm ürünleri belli bir sıraya göre görülürken, SPSS'te her ürün ayrı ayrı analiz edilip ilgili değerlere ulaşılmaktadır. Burada amaç, bütüncül/genel ve ayrı ayrı verilen değerleri karşılaştırmak ve katılımcıların tek bir ürün değil tüm ürünler hakkındaki genel ve spesifik görüşlerini harmanlamaktır. Yani hem tek hem de bütüncül bir analiz ortaya koymaktır.
2.	Programa İlişkin Ölçüm Raporu	Programların sınıf bazında genel aritmetik ortalama değerlerinin yorumu SPSS programıyla yapılmıştır. Rasch'ta ise programların nitelik sıralaması yapılmıştır. Dolayısıyla "Ölçeğin Tüm Maddelere İlişkin Genel Aritmetik Ortalaması" değerinin sınıf bazında incelenmesi sonucu ile programların nitelik sıralaması karşılaştırılmıştır.	Rasch'ta değerlendirilen birden fazla ürünün analizi sonucunda ortaya konan etkililik sıralamasının yanında bu ürünlere ilişkin katılım düzeylerine de gerek duyulmaktadır. Bu gereklilik SPSS Programını kullanma gereksinimini doğurmuştur. Başka bir deyişle, Rasch en etkili olandan en etkisiz olana doğru bir sıralama yapılmakta ancak katılım düzeyi verilmemektedir. SPSS ise bu sıralamanın katılım düzeyi/dercesini yani aritmetik ortalamasını sunmaktadır. Diğer yandan sadece SPSS Programı, anket ürün kadar uygulama sonucu verebilir.
3.	Jürilerin katılım/cömertlik karşılaştırmaları	Cinsiyet değişkeni-p değeri Rasch modelinde, cinsiyet değişkeni gözetilmeksizin bir sıralama yapılmıştır. Sıralama karışık olup olumlu durumun hangi cinsiyet lehine olduğunu belirtmemiştir. SPSS bu boşluğu doldurarak anlamlı farklılığın ve olumlu durumun kimin lehine olduğunu belirtmiştir.	Jürilerin (Katılımcı) en katı (puanı az) olandan en cömert (puanı fazla) olana doğru cinsiyet gözetilmeksizin bir sıralama yapılmıştır (Rasch Ölçme Modeli). Ancak katılım düzeyinin yönü ve olumlu-olumsuzluk boyutu verilmemiştir. SPSS programında jürilerin hem cinsiyet hem de katılım düzeyinin yönü ve anlamlı farklılık durumu, faktör bazında analiz edilerek daha özelliikli bir analize ihtiyaç duyulmuştur.
4.	Yeterlik İstatistikleri	SPSS'te her maddenin aritmetik ortalamasının karşılaştırılması yapılmaktadır. Rasch modelinde ise maddelerin genel olarak hangi değerler arasında olup en yüksek ve en düşük olanlar belirtilerek tüm sonuçlar karşılaştırılmaktadır.	Rasch'ta madde güçlükleri arasında anlamlı farklılık temel unsur olmakla birlikte maddelerin en zor olandan en kolay olana doğru bir sıralama yapılır. Ancak katılım düzeyinin, faktör bazında genel aritmetik ortalaması ve tüm maddelerin toplam aritmetik ortalaması SPSS programı analizleriyle gerçekleşmektedir. Özellikle madde güçlükleri arasındaki anlamlı farklılık (Rasch) ve katılım düzeylerini (SPSS) birlikte analiz etme amacı, iki analiz programını gerekli kılmıştır.

Maxqda Programı: Katılımcıların ürün olarak ifade edilen "programlar" hakkındaki görüşlerinin analiz boyutunu temsil etmektedir.